A DESCENT METHOD FOR VARIATIONAL INEQUALITIES IN HILBERT SPACES

R. Sadhu[1], C. Nahak[2]

In this article, a variational inequality problem is reformulated as an optimization problem. An adaptive descent method for solving the reformulated optimization problem is proposed. The method is structured by the introduction of a generalised class of differentiable gap function. Global convergence of the proposed descent method is also proved in a Hilbert space.

Keywords: Variational inequality, Gap function, Optimization, Descent method, Global convergence

1. Introduction

Let H be a separable Hilbert space and $\mathbb{S}$ be a nonempty closed convex subset of H, $\langle .,. \rangle$ denotes the standard inner product on H and $\|.\|$ represents the usual norm induced by the inner product i.e., $\|x\| = \sqrt{\langle x, x \rangle}$ for all $x \in H$. Let $F : H \to H$ be a mapping, then the variational inequality problem (in short,VIP) is to find $x^* \in \mathbb{S}$ such that,

$$(VIP) \ \langle F(x^*), x - x^* \rangle \geq 0, \ \forall x \in \mathbb{S} \tag{1}$$

The applicability of VIP in diverse areas such as partial differential equation, operation research and mathematical economics, has craved the attention of several researchers and various solution methods for VIP have been developed. For a comprehensive survey, readers are referred to Harker and Pang [5] and the references therein.

In the last two decades a considerable amount of research has been devoted in reformulating a VIP as an equivalent optimization problem. The methodology involves a non-negative real-valued functions g on the search space $\mathbb{S}$, with the property $g(x) = 0$ if and only if x is a solution to the VIP. These functions allows to cast the VIP as the following minimization problem:

$$minimize \ g(x) \ \text{subject to} \ x \in \mathbb{S}. \tag{2}$$

Such a function g is often called as gap function and was first introduced by Auslenders [7]. Although Auslenders gap function was not differentiable but it possessed some excellent global properties, the differentiable gap function was first proposed by Auchmuty [8] and Fukushima [9]. By using Fukushima's regularized gap function the VIP was successfully reformulated as a differentiable optimization problem. Various extensions and modifications of these differentiable gap functions were accomplished by many authors; e.g., see ([10], Chapter 4) and references therein. Such merit functions and their associated descent algorithm in finite dimension, comprises a huge portion of the literature; e.g., see [14, 15, 16, 17, 18] Most of these descent type methods have been restricted to the finite dimensional space so far, a quest for the applicability and convergence of these methods in the infinite dimensional

---

[1]Department of Mathematics, Indian Institute Of Technology, Kharagpur, India, e-mail: rudrajitsadhu@gmail.com

[2]Department of Mathematics, Indian Institute Of Technology, Kharagpur, India, e-mail: cnahak@maths.iitkgp.ac.in

setting encourages many authors. Fukushima's regularized gap function was extended to the Hilbert space setting and a descent framework was proposed by Konnov, Kum and Lee see, ([11]). The global convergence of the method was also proved under the strong monotonicity assumption on the cost mapping. Note that a recent work due to Zhu and Marcotte [12] deals with a generalization of Fukushima's method in Banach spaces. Using Auchmutys gap function, Noor presented fixed-point algorithms for VIPs under a Hilbert space setting; see [13]. For more recent iterative methods in VIP see, [1, 2, 3, 4, 6]

In this paper, we propose a generalized class of differentiable gap function for solving the VIP in a Hilbert space. This class of gap function generalizes the regularized gap function proposed by Fukushima. Here we extensively study different properties of the proposed gap function and use it to develop an adaptive descent method. The global convergence of the scheme is also proved under some mild assumption on the cost mapping.

The paper is organized as follows. Some basic definition and results are recalled in Section 2. Section 3 addresses the formulation of the general class of gap function. Section 4 discusses some properties of the proposed gap function. Section 5 defines the auxiliary problem. Section 6 deals with the generalised descent algorithm and its global convergence are proved in Section 7. A brief conclusion of our work is presented in Section 9.

2. Preliminaries

In this section, we will recall some basic definition and properties that will be useful in the subsequent sections.

Definition 2.1. *An operator $F : \mathbb{S} \subseteq H \to H$ is said to be monotone on $\mathbb{S}$ if*

$$\langle F(x) - F(y), x - y \rangle \geq 0, \ \forall x, y \in \mathbb{S} \tag{3}$$

F is said to be strictly monotone if strict inequality holds in (3).

Definition 2.2. *F is said to be strongly monotone with modulus μ on $\mathbb{S}$ if it satisfies,*

$$\langle F(x) - F(y), x - y \rangle \geq \mu ||x - y||^2, \ \forall x, y \in \mathbb{S} \tag{4}$$

for some $\mu > 0$.

Definition 2.3. *F is said to be strongly pseudo-monotone with modulus η on $\mathbb{S}$ if for every $x, y \in \mathbb{S}$ we have,*

$$\langle F(x), y - x \rangle \geq 0 \Rightarrow \langle F(y), y - x \rangle \geq \eta ||x - y||^2 \tag{5}$$

Definition 2.4. *F is said to be coercive on $\mathbb{S}$ if there exist a $x^o \in \mathbb{S}$ such that*

$$\lim_{x \in \mathbb{S}, ||x|| \to \infty} \frac{\langle F(x), x - x^o \rangle}{||x - x^o||} = \infty. \tag{6}$$

It is to be noted that, every strongly monotone function on $\mathbb{S}$ is coercive on $\mathbb{S}$.

Definition 2.5. *Let $(X, ||.||_X)$ and $(Y, ||.||_Y)$ be Banach spaces, and U be an open subset of X. A function $F : U \to Y$ is called Fréchet differentiable at $x \in U$ if there exists a bounded linear operator $A : X \to Y$ such that*

$$\lim_{u \to 0} \frac{||F(x + u) - F(x) - Au||_Y}{||u||_X} = 0 \tag{7}$$

Definition 2.6. *A function $f : \mathbb{S} \to \mathbb{R}$ is said to be strongly convex on $\mathbb{S}$ with parameter $a > 0$, if for every $x, y \in \mathbb{S}$ and $\lambda \in [0, 1]$ we have,*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{a\lambda(1 - \lambda)}{2}||x - y||^2. \tag{8}$$

Moreover if f isFréchet differentiable, strong convexity of f can also be defined as,

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \xi ||x - y||^2, \ \forall x, y \in \mathbb{S} \tag{9}$$

for some $\xi > 0$. It is hence worth noting that strong convexity of f ensures $\nabla f(x)$ is strongly monotone.

Definition 2.7. [11] *[Gap function] A function $\phi : \mathbb{S} \to \mathbb{R} \cup \{+\infty\}$ is called a gap function for the variational inequality problem* (1) *if it satisfies,*

(i) $\phi(x) \geq 0, \ \forall x \in \mathbb{S}$

(ii) $\phi(x) = 0$ *if and only if x solves the (VIP).*

3. Extended gap function formulation

Let $\Omega(x,y) : \mathbb{S} \times \mathbb{S} \to \mathbb{R}$ be a mapping which satisfies the following property,

- $\Omega(x,y) \geq 0, \ \forall x, y \in \mathbb{S}$
- Ω is continuously differentiable on $\mathbb{S} \times \mathbb{S}$
- $\Omega(x,y)$ is strongly convex on $\mathbb{S}$ with respect to 'y' for all $x \in \mathbb{S}$
- $\Omega(x,x) = 0, \ \forall x \in \mathbb{S}$
- $\nabla_y \Omega(x,x) = 0, \ \forall x \in \mathbb{S}$
- $\nabla_y \Omega(x,y)$ is Lipschitz continuous in the variable 'y', on any bounded subset of $\mathbb{S}$.

Some examples of $\Omega(x,y)$ that satisfies the above conditions are given below,

$\Omega(x,y) = \frac{1}{2} \|x - y\|^2$,

$\Omega(x,y) = \frac{1}{2} \langle M(x-y), x-y \rangle$,

$\Omega(x,y) = \frac{1}{2} \langle B_x(x-y), x-y \rangle$,

where M and B_x are symmetric and positive definite operator on a real Hilbert space H. Note that such a bi-function $\Omega(x,y)$, defined above, is a proximity measure between x and y. Let us now define,

$$h(x,y) = \langle F(x), x-y \rangle - \Omega(x,y) \tag{10}$$

$$g(x) = \max_{y \in \mathbb{S}} h(x,y) = h(x, Tx) \tag{11}$$

where the maximizer Tx is unambiguously defined, since h is the sum of a linear and a strongly concave term. Thus

$$g(x) = \langle F(x), x - Tx \rangle - \Omega(x, Tx). \tag{12}$$

On account of Tx being a solution to the maximization problem (11) it satisfies the following variational inequality,

$$\langle F(x) + \nabla_y \Omega(x, Tx), z - Tx \rangle \geq 0, \ \forall z \in \mathbb{S} \tag{13}$$

Theorem 3.1. *The function g defined as in* (12) *is a gap function.*

Proof. Let $x \in \mathbb{S}$, by (10) $h(x,x) = 0$ and thus by (11) $g(x) \geq h(x,x) = 0$. So, $g(x) \geq 0, \forall \ x \in \mathbb{S}$. Now, let us assume that x solves the VIP. So, $\langle F(x), y - x \rangle \geq 0 \forall y \in \mathbb{S}$ and so, $h(x,y) = \langle F(x), x-y \rangle - \Omega(x,y) \leq 0, \ \forall y \in \mathbb{S}$. Hence, $g(x) = \max_{y \in \mathbb{S}} h(x,y) \leq 0$. This implies $g(x) = 0$

Conversely, let $g(x) = 0$. Thus, $h(x,y) \leq 0, \ \forall y \in \mathbb{S}$ and hence x is a solution to the optimization problem $\max_{y \in \mathbb{S}} h(x,y)$ (as $h(x,x) = 0$). Therefore x satisfies the variational inequality,

$$\langle \nabla_y h(x,x), z - x \rangle \leq 0, \ \forall z \in \mathbb{S}$$
$$\langle F(x) + \nabla_y \Omega(x,x), z - x \rangle \geq 0, \ \forall z \in \mathbb{S}$$
$$\langle F(x), z - x \rangle \geq 0, \ \forall z \in \mathbb{S}$$

Thus x solves the VIP. $\square$

The above theorem ensures that the function g is a gap function. We now provide a fixed point characterization of the solution of VIP in terms of the mapping T

Theorem 3.2. *x is a solution to the VIP if and only if x is a fixed point of the mapping T, i.e. $x = Tx$.*

Proof. Let us assume $Tx = x$, then $g(x) = 0$ and x solves the VIP.

Conversely, let x solves the VIP. So, $\langle F(x), y - x \rangle \geq 0, \ \forall y \in \mathbb{S}$

Since $Tx \in \mathbb{S}, \ \forall x \in H, \ \langle F(x), Tx - x \rangle \geq 0$. Since Tx is a solution of the optimization problem $\max\limits_{y \in \mathbb{S}} h(x, y)$, it satisfies the variational inequality

$$\langle \nabla_y h(x, Tx), y - Tx \rangle \leq 0, \ \forall y \in \mathbb{S}$$

$$\langle F(x) + \nabla_y \Omega(x, Tx), x - Tx \rangle \geq 0$$

$$\langle \nabla_y \Omega(x, Tx), x - Tx \rangle \geq \langle F(x), Tx - x \rangle \geq 0 \tag{14}$$

Therefore $\langle \nabla_y \Omega(x, Tx), Tx - x \rangle \leq 0$. Since, $\Omega(x, y)$ is strongly convex in y, there exist a schalar $\xi > 0$ which satisfies

$$\langle \nabla_y \Omega(x, Tx) - \nabla_y \Omega(x, x), Tx - x \rangle \geq \xi \|Tx - x\|^2 \geq 0 \tag{15}$$

Thus by (14) and (15) $\|Tx - x\| = 0$ i.e., $Tx = x$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

4. Properties of gap function

Let us first prove that the gap function g is differentiable on $\mathbb{S}$. We begin with a technical lemma.

Lemma 4.1. *([11] ,Lemma 3.1) Let $h : H \times \mathbb{S} \to \mathbb{R}$ be a function such that $\nabla_x h(x, y)$ exists and is continuous on $H \times \mathbb{S}$. Define two function as follows*

$$\Psi x = \max_{y \in \mathbb{S}} h(x, y), \ \forall x \in \mathbb{S}$$

$$Mx = \{y \in \mathbb{S} \mid h(x, y) = \Psi x\}$$

Assume that Mx is a single tone for all $x \in H$, and M is a continuous function on H, then Ψ is continuously differentiable and the gradient of Ψ is given by

$$\nabla \Psi x = \nabla_x h(x, Mx)$$

Hence onwards in this chapter, by 'differentiability' we will mean 'Frechet-differentiability.'

Theorem 4.1. *The gap-function g defined by (12) is continuous (resp. continuously differentiable) whenever F is continuous (resp. continuously differentiable)on $\mathbb{S}$. In particular if g is continuously differentiable, the gradient of g is given by*

$$\nabla g(x) = F(x) - \langle \nabla F(x)(.), Tx - x \rangle - \nabla_x \Omega(x, Tx).$$

Proof. By continuity of F and Ω we have $h(x, y)$ defined by (10) is continuous. By definition ,

$$g(x) = \max_{y \in \mathbb{S}} h(x, y) = h(x, Tx)$$

Let $x \in H$ and $\epsilon > 0$ be arbitrary, let $z = x + h$, for some $h \in H$ by (11)

$$g(z) = h(z, Tz) \leq h(x, Tz) + \epsilon \ [\text{by continuity of } h(x, y)]$$

$$\leq g(x) + \epsilon$$

whenever $\|z - x\| \leq \delta$, for some $\delta > 0$ depending upon ϵ . Since we may similarly obtain the same inequality with x and z switched, continuity of g follows.

Suppose that F is continuously differentiable. Define $h : H \times \mathbb{S} \to \mathbb{R}$ by,

$$h(x, y) = \langle F(x), x - y \rangle - \Omega(x, y).$$

Then we apply Lemma 4.1 to the case $\Psi x = \max\limits_{y \in \mathbb{S}} h(x, y)$ and $Mx = Tx$ so that we can obtain

$$\nabla \Psi x = \nabla_x h(x, Mx) = F(x) - \langle \nabla F(x)(.), Tx - x \rangle - \nabla_x \Omega(x, Tx)$$

Thus we have,

$$\nabla g(x) = F(x) - \langle \nabla F(x)(.), Tx - x \rangle - \nabla_x \Omega(x, Tx).$$

This completes the proof. $\qquad\qquad\square$

Proposition 4.1. [11]*) Let F be continuously differentiable and the gradient $\nabla F(x)$ be positive (or strictly monotone) for all $x \in \mathbb{S}$, i.e., $\langle \nabla F(x)d, d \rangle > 0$ for all $d(\neq 0) \in H$. If x is a stationary point of g on $\mathbb{S}$,i.e. $\langle \nabla g(x), y - x \rangle \geq 0$, $\forall y \in \mathbb{S}$ then x is a global optimal solution of* (2) *and hence solves the VIP.*

The following property of g is essential for constructing a descent method for minimizing g over $\mathbb{S}$.

Theorem 4.2. *Let x^* be a solution of the VIP. If F is strongly pseudo monotone with modulus $\mu > 0$ on $\mathbb{S}$ and the gradient of Ω with respect to y is Lipschitz continuous with modulus $L_{\nabla_y \Omega_y}$ on $\mathbb{S}$ then there exists a positive scalar α such that*

$$g(x) \geq \alpha ||x - x^*||^2, \ \forall x \in \mathbb{S}$$

Proof. Since x^* is a solution to the VIP, so $\langle F(x^*), x - x^* \rangle \geq 0$, $\forall x \in \mathbb{S}$. By pseudo-monotonicity of F we have,

$$\langle F(x), x - x^* \rangle \geq \mu ||x - x^*||^2, \ \forall x \in \mathbb{S}$$

Let $x_t = x + t(x^* - x)$, $t \in [0, 1]$.
Now by convexity of $\Omega(z, y)$ and Lipschitz continuity of $\nabla_y \Omega(z, y)$ with respect to y, we get

$$\begin{aligned}
\Omega(z, x_t) - \Omega(z, x) &\leq \langle \nabla_y \Omega(z, x_t), x_t - x \rangle \\
&\leq \langle \nabla_y \Omega(z, x_t) - \nabla_y \Omega(z, x), x_t - x \rangle \\
&\leq ||\nabla_y \Omega(z, x_t) - \nabla_y \Omega(z, x)||.||x_t - x|| \\
&\leq L_{\nabla_y \Omega_y} ||x_t - x||^2.
\end{aligned}$$

Thus we have, $g(x) \geq \langle F(x), x - y \rangle - \Omega(x, y)$, $\forall y \in \mathbb{S}$.
Let $y = x_t \in \mathbb{S}$. Therefore,

$$\begin{aligned}
g(x) &\geq \langle F(x), x - x_t \rangle - \Omega(x, x_t) \\
&\geq t \langle F(x), x - x* \rangle - (\Omega(x, x_t) - \Omega(x, x)) \\
&\geq t\mu ||x - x^*||^2 - L_{\nabla_y \Omega_y} ||x_t - x||^2 \\
&= (t\mu - L_{\nabla_y \Omega_y} t^2) ||x - x^*||^2 \\
&= t L_{\nabla_y \Omega_y} \left(\frac{\mu}{L_{\nabla_y \Omega_y}} - t \right) ||x - x^*||^2
\end{aligned}$$

Let $t = \min\{1, \frac{\mu}{2 L_{\nabla_y \Omega_y}}\}$; we chose,

$$\alpha = \begin{cases} \mu - L_{\nabla_y \Omega_y} & \text{if } \mu > 2 L_{\nabla_y \Omega_y}, \\ \frac{\mu^2}{4 L_{\nabla_y \Omega_y}} & \text{if } \mu \leq 2 L_{\nabla_y \Omega_y}. \end{cases}$$

Thus $\alpha > 0$ and $g(x) \geq \alpha ||x - x^*||^2$ for all $x \in \mathbb{S}$ $\qquad\qquad\square$

Remarks: The above inequality infers that g is non-negative and has a bounded level set.

5. Auxiliary variational inequality

Let $\Gamma(x,y) : \mathbb{S} \times \mathbb{S} \to H$ be a continuous mapping (the auxiliary mapping), which is strongly monotone with respect to the variable y. The auxiliary variational inequality problem is stated as: for each fixed $x \in \mathbb{S}$, find a point $w \in \mathbb{S}$ which satisfies the following inequality,

$$(AVIP(x)) \ \langle \Gamma(x,w) - \Gamma(x,x) + F(x), y - w \rangle \geq 0, \ \forall y \in \mathbb{S}. \tag{16}$$

Due to the strong monotonicity of $\Gamma(x,y)$ with respect to y, for every fixed $x \in S$ there exist a unique solution to the above auxiliary variational inequality problem. Let us now construct a function $W : \mathbb{S} \to \mathbb{S}$, such that $W(x)$ is the unique solution to $(AVIP(x))$ for each point $x \in S$ i.e.,

$$\langle \Gamma(x, W(x)) - \Gamma(x,x) + F(x), y - W(x) \rangle \geq 0, \ \forall y \in \mathbb{S}. \tag{17}$$

Theorem 5.1. *x is a solution to the VIP if and only if it is a fixed point of the mapping W i.e., $W(x) = x$.*

Proof. Let us first suppose that x is a fixed point of W i.e., $W(x) = x$. Thus substituting x in place of $W(x)$ in (17) we obtain,

$$\langle F(x), y - x \rangle \geq 0, \ \forall y \in \mathbb{S}$$

and hence x solves the VIP (1).
Conversely, suppose that x is a solution to the (VIP), then,

$$\langle F(x), W(x) - x \rangle \geq 0$$

and it follows from (17) that

$$\langle \Gamma(x, W(x)) - \Gamma(x,x) + F(x), x - W(x) \rangle \geq 0$$

Adding the two preceding inequalities, we obtain

$$\langle \Gamma(x, W(x)) - \Gamma(x,x), x - W(x) \rangle \geq 0 \tag{18}$$

Now by strong monotonicity of $\Gamma(x.y)$ with respect to y we have

$$\langle \Gamma(x, W(x)) - \Gamma(x,x), W(x) - x \rangle \geq \theta ||W(x) - x||^2 \geq 0 \tag{19}$$

where θ is the modulus of strong monotonicity of Γ. Thus by comparing (18) and (19), we get $W(x) = x$. $\qquad\square$

6. Generalised descent algorithm

Theorem 5.1 suggests a fixed-point algorithm for solving a VIP. For given $x^k \in \mathbb{S}$ consider the auxiliary variational inequality problem: Find $W(x^k) \in \mathbb{S}$ such that

$$(AVIP(x^k)) \ \langle \Gamma(x^k, W(x^k)) - \Gamma(x^k, x^k) + F(x^k), y - W(x^k) \rangle \geq 0, \ \forall y \in \mathbb{S} \tag{20}$$

in our approach, instead of choosing $x^{k+1} = W(x^k)$, we will use

$$d^k = W(x^k) - x^k$$

as a descent direction for the merit function g at x^k and incorporate Armijo type line search technique to find the next iterate x^{k+1}. If $x^k = W(x^k)$, the solution to VIP is obtained, otherwise, the vector $d^k = W(x^k) - x^k$, is a descent direction for g at x^k, under conditions to be imposed latter.

The liberty to choose Γ often makes the problem AVIP(x^k), easier to solve than the original VIP. In this chapter, we impose a specific structure to the auxiliary function Γ and presents a general descent scheme. Let us first assert the generic algorithm.

General Algorithm

Step 0: Chose, $x^0 \in \mathbb{S}$, set tolerance factor $\epsilon > 0$ and select parameters γ, β and σ from the interval $(0, 1)$. Set $k = 0$.

Step 1: At k^{th} iteration of the algorithm, compute $W(x^k)$ by solving the auxiliary variational inequality problem AVIP(x^k). Set $d^k = W(x^k) - x^k$

Step 2: *Line search* if $g(x^k + d^k) < \gamma g(x^k)$, set $x^{k+1} = x^k + d^k$.
Otherwise select the smallest positive integer m such that
$g(x^k) - g(x^k + \beta^m d^k) \geq -\sigma \beta^m ||d^k||^2$; set $\alpha_k = \beta^m$,
and set $x^{k+1} = x^k + \alpha_k d^k$.

Step 3: STOP if $||x^{k+1} - x^k|| < \epsilon$. Otherwise, increase k by one and return to Step 1.

The flexibility, to choose the regularizing term Ω in g, enables us to construct a gap function which suits the problem most. The function Ω should be chosen in such a fashion that g and its gradient map ∇g are easily computable at any point x in H. The gap function g acts as a merit function and thus monitors the iterates to a solution of the VIP. By sincere choice of Γ and Ω one can thus construct an excellent algorithm that will guide the sequence generated by it, to a solution of the VIP. In the following sections, we introduce the specific choice of Γ and the associated convergence frameworks.

7. Descent method and its convergence

In this section, Ω assumes the general form introduced in Section 3, and the auxiliary mapping is chosen as
$$\Gamma(x,y) = \nabla_y \Omega(x,y).$$
For this specific choice of Γ, we observe that,

$$h(x,y) = \langle F(x), x - y \rangle - \Omega(x,y) \tag{21}$$

$$T(x) = \text{argmax}_{y \in \mathbb{S}} h(x,y) = W(x) \tag{22}$$

$$g(x) = \max_{y \in \mathbb{S}} h(x,y) = h(x, Wx)$$

$$g(x) = \langle F(x), x - W(x) \rangle - \Omega(x, W(x)) \tag{23}$$

where $W(x)$ is the unique solution of the auxiliary variational inequality $(AVIP(x))$. From (gradient function), we have

$$\nabla g(x) = F(x) - \langle \nabla F(x)(.), W(x) - x \rangle - \nabla_x \Omega(x, W(x)). \tag{24}$$

Following the analysis above, we introduce the general descent framework for VIP.

Algorithm

Step 0: Chose, $x^0 \in \mathbb{S}$, set tolerance factor $\epsilon > 0$ and select parameters γ, β *and* σ from the interval $(0,1)$. Set $k = 0$.

Step 1: At k^{th} iteration of the algorithm, compute $W(x^k)$ by solving the auxiliary variational inequality problem
$$AVIP(x^k) \ \langle \nabla_y \Omega(w, x^k) + F(x^k), y - w \rangle \geq 0, \ \forall y \in \mathbb{S}$$
and let $d^k = W(x^k) - x^k$.

Step 2: *Line search* if $g(x^k + d^k) < \gamma g(x^k)$, set $x^{k+1} = x^k + d^k$.
Otherwise select the smallest positive integer m such that
$$g(x^k) - g(x^k + \beta^m d^k) \geq -\sigma \beta^m ||d^k||^2; \tag{25}$$
set $\alpha_k = \beta^m$,
and set $x^{k+1} = x^k + \alpha_k d^k$.

Step 3: STOP if $||x^{k+1} - x^k|| < \epsilon$.Otherwise, increase k by one and return to Step 1.

The following assumptions are essential for global convergence of the algorithm.

Assumption 7.1. *In addition to the condition stated in Sec.3, let the function $\Omega(x, y)$ also satisfies ,*

- *$\langle \nabla_x \Omega(x, y) + \nabla_y \Omega(x, y), x - y \rangle \geq 0, \ \forall x, y \in \mathbb{S}$*
- *$\nabla_y \Omega(x, y)$ is Lipschitz continuous in x on any bounded subset of $\mathbb{S}$.*
- *$\nabla_x \Omega(x, y)$ is Lipschitz continuous on any bounded subset of $\mathbb{S} \times \mathbb{S}$.*

Theorem 7.1. *If F is strongly monotone with modulus μ on $\mathbb{S}$, Ω satisfies assumption and x^k is not a solution of VIP, then $d^k = W(x^k) - x^k$ satisfies*

$$\langle \nabla g(x^k), d^k \rangle < -\eta ||d^k||^2 \tag{26}$$

for some $\eta > 0$, i.e. d^k is a feasible descent direction of g at x^k.

Proof. Let us temporarily ignore the iteration index k for notational simplicity. Since $d = W(x^k) - x^k$, we have from (24) that

$$
\begin{aligned}
\langle \nabla g(x), d \rangle &= \langle F(x), W(x) - x \rangle - \langle \nabla F(x)(W(x) - x), W(x) - x \rangle \\
&\quad - \langle \nabla_x \Omega(x, W(x)), W(x) - x \rangle \\
&= \langle \nabla_y \Omega(x, W(x)) + F(x), W(x) - x \rangle \\
&\quad - \langle \nabla_x \Omega(x, W(x)) + \nabla_y \Omega(x, W(x)), W(x) - x \rangle \\
&\quad - \langle \nabla F(x)d, d \rangle
\end{aligned}
\tag{27}
$$

Since $W(x)$ is a solution of AVIP(x), the first term of (27) is non-positive. From the above assumption, the second term of (27) is also non-positive. By strong monotonicity of F ,

$$\langle \nabla F(x)d, d \rangle \geq \mu ||d||^2.$$

Letting $\eta = \mu$ the result follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The above theorem ensures that the direction d^k is in fact a descent direction for g at x^k. Next, we present some lemmas which will play an important role to prove the global convergence of the algorithm.

Lemma 7.1. *The mapping $W : \mathbb{S} \to \mathbb{S}$ is Lipschitz continuous on any bounded subset of $\mathbb{S}$*

Proof. By equation (21),

$$g(x) = \max_{y \in \mathbb{S}} h(x, y) = h(x, Wx)$$

Hence Wx satisfies,

$$\langle \nabla_y h(x, Wx), z - Wx \rangle \leq 0, \ \forall z \in \mathbb{S} \tag{28}$$

Let $x_1, x_2 \in H$ then by (28) we have,

$$\langle \nabla_y h(x_1, Wx_1), z - Wx_1 \rangle \leq 0, \ \forall z \in \mathbb{S} \tag{29}$$

$$\langle \nabla_y h(x_2, Wx_2), z' - Wx_2 \rangle \leq 0, \ \forall z' \in \mathbb{S} \tag{30}$$

Substituting $z = Wx_2$ in (29) and $z' = Wx_1$ in (30), and adding we obtain

$$\langle \nabla_y h(x_1, Wx_1) - \nabla_y h(x_2, Wx_2), Wx_1 - Wx_2 \rangle \geq 0 \tag{31}$$

or

$$\langle \nabla_y h(x_1, Wx_1) - \nabla_y h(x_2, Wx_2) + \nabla_y h(x_2, Wx_1) - \nabla_y h(x_2, Wx_1), Wx_1 - Wx_2 \rangle \geq 0 \tag{32}$$

Using the strong monotonicity of F and the strong convexity of Ω in the second variable, (32) yields

$$\langle \nabla_y h(x_1, Wx_1) - \nabla_y h(x_2, Wx_1), Wx_1 - Wx_2 \rangle$$
$$\geq \langle \nabla_y h(x_2, Wx_2) - \nabla_y h(x_2, Wx_1), Wx_1 - Wx_2 \rangle$$
$$= \langle \nabla_y \Omega(x_2, Wx_1) - \nabla_y \Omega(x_2, Wx_2), Wx_1 - Wx_2 \rangle$$
$$+ \langle F(x_1) - F(x_2), Wx_1 - Wx_2 \rangle$$
$$\geq l||Wx_1 - Wx_2||^2$$

for some positive constant l.

By virtue of Schwarz's inequality, the above inequality gives

$$l||Wx_1 - Wx_2||^2 \leq ||\nabla_y h(x_1, Wx_1) - \nabla_y h(x_2, Wx_1)|| \, ||Wx_1 - Wx_2||$$

or

$$||Wx_1 - Wx_2|| \leq \frac{1}{l}||\nabla_y h(x_1, Wx_1) - \nabla_y h(x_2, Wx_1)||$$
$$\leq \frac{1}{l}(||F(x_1) - F(x_2)|| + ||\nabla_y \Omega(x_1, Wx_1) - \nabla_y \Omega(x_2, Wx_1)||)$$

Thus the Lipschitz continuity of F and that of $\nabla_y \Omega(x, y)$ in the variable x on any bounded subset of $\mathbb{S}$ enforces W to be Lipschitz continuous. $\square$

Proposition 7.1. *([11] , Lemma 4.2) Let $f : H \to \mathbb{R}$ be a mapping, such that the gradient function ∇f is Lipschitz continuous on a convex set $D \subseteq H$ with modulus $L,0$ then for any $x, y \in D$ and $t \in \mathbb{R}$, we have,*

$$f(x + t(y - x)) \leq f(x) + t\langle \nabla f(x), y - x \rangle - \frac{1}{2}Lt^2||x - y||^2 \tag{33}$$

Lemma 7.2. *∇g is Lipschitz continuous on any bounded subset D of $\mathbb{S}$*

Proof. Let $x_1, x_2 \in D$, we have

$$||\nabla g(x_1) - \nabla g(x_2)|| = ||F(x_1) - F(x_2)$$
$$- [(\nabla F(x_1))^*(W(x_1) - x_1) - (\nabla F(x_2))^*(W(x_2) - x_2)]$$
$$- (\nabla_x \Omega(x_1, W(x_1)) - \nabla_x \Omega(x_2, W(xx_2)))||$$
$$\leq ||F(x_1) - F(x_2)||$$
$$+ ||(\nabla F(x_1))^*(W(x_1) - x_1) - (\nabla F(x_2))^*(W(x_2) - x_2)||$$
$$+ ||\nabla_x \Omega(x_1, W(x_1)) - \nabla_x \Omega(x_2, W(xx_2))||$$
$$\leq ||F(x_1) - F(x_2)||$$
$$+ ||[(\nabla F(x_1))^* - (\nabla F(x_2))^*](W(x_1) - x_1)||$$
$$+ ||(\nabla F(x_2))^*[W(x_1) - W(x_2) + x_2 - x_1]||$$
$$+ ||\nabla_x \Omega(x_1, W(x_1)) - \nabla_x \Omega(x_2, W(xx_2))||$$
$$\leq L_F||x_1 - x_2|| + L_{\nabla F}||x_1 - x_2|| \, ||W(x_1) - x_1||$$
$$+ (1 + \theta)||(\nabla F(x_2))^*|| \, ||x_1 - x_2||$$
$$+ L_{\nabla_x \Omega}(1 + \theta)||x_1 - x_2||$$
$$= [L_F + L_{\nabla F}||W(x_1) - x_1|| + (1 + \theta)$$
$$(||(\nabla F(x_2))^*|| + L_{\nabla_x \Omega})]||x_1 - x_2||$$

where L_F, $L_{\nabla F}$, θ and $L_{\nabla_x \Omega}$ are the Lipschitz constants for F ,∇F, W and $\nabla_x \Omega$ on D, respectively. Since W, F and ∇F are Lipschitz continuous and D is bounded, there exist a

constant $L < \infty$ such that

$$L_F + L_{\nabla F}||W(x_1) - x_1|| + (1 + \theta)(||(\nabla F(x_2))^*|| + L_{\nabla_x \Omega}) < L,$$

that is, ∇g is Lipschitz continuous on D. $\qquad\square$

We are now in a position to state the global convergence result in the Hilbert space.

Theorem 7.2 (Global Convergence). *Let $\{x^k\}$ be a sequence generated by the iteration $x^{k+1} = x^k + \alpha_k d^k$; $k = 0, 1, 2, ...,$ where d^k is given by $d^k = W(x^k) - x^k$ and $\alpha_k \in [0, 1]$; is determined by the Armijo-type steplength rule (25). Assume that $F : H \to H$ is differentiable and strongly monotone with modulus $\mu > 0$ on $\mathbb{S}$. Further assume that F and ∇F are Lipschitz continuous on each bounded subset of $\mathbb{S}$. Let Ω be defined as in sec3 and satisfies assumption 7.1. Then whenever the positive constant σ in the Armijo-type steplength rule (25) is chosen sufficiently small so that $\sigma < \mu$, then the generated sequence $\{x^k\}$ lies in $\mathbb{S}$ and converges to a unique solution of (VIP)(1) for any starting point $x^0 \in \mathbb{S}$.*

Proof. To prove the theorem it is enough to follow the subsequent steps.
(i) *The sequence $\{x^k\}$ lies in $\mathbb{S}$ and is bounced.*
    Since x^k and $x^k + d^k$ both belong to $\mathbb{S}$ and since $0 < \alpha_k < 1$, it follows from the convexity of $\mathbb{S}$ that the sequence $\{x^k\}$ lies in $\mathbb{S}$. Now by Theorem 7.1 and Step 2 (line search rule) of the algorithm, the sequence $\{g(x^k)\}$ is strictly decreasing. Thus by Theorem 4.2, $\{x^k\}$ is bounded.

(ii) *The sequence $\{g(x^k)\}$ converges to 0 as k tends to infinity.*

Case 1: if $g(x^k + d^k) \leq \gamma g(x^k)$, for $\gamma < 1$, holds infinitely often. Then there exist a subsequence $\{x^k\}$, $k \in \mathcal{K}$, such that

$$\lim_{k \to \infty} g(x^k) = 0 \ k \in \mathcal{K}.$$

Case 2: Let us now consider the case where $g(x^k + d^k) \leq \gamma g(x^k)$, holds for only finitely many values of the index k; i.e. there exist index $\bar{K}_1$ such that

$$g(x^k + d^k) > \gamma g(x^k), \ k \geq \bar{K}_1.$$

Then for $k \geq \bar{K}_1$, by Armijo line search rule we have,

$$g(x^{k+1}) - g(x^k) \leq \sigma \alpha_k ||d^k||^2 e \tag{34}$$

thus by (34) we observe that the sequence $\{g(x^k)\}$ for $k \geq \bar{K}_1$ is monotonically decreasing and by Theorem 3.1 it is bounded below. By taking limit on both side of (34) we get,

$$\lim_{k \longrightarrow \infty} \alpha_k ||d^k|| = 0. \tag{35}$$

Claim 1: $\alpha_k \geq \alpha' > 0$ for all $k \geq \bar{K}_1$.

As W is Lipschitz continuous $\{x^k\}$ is bounded (by (i)) we can easily show that $\{d^k\}$ is also bounded. Therefore, there exists a closed, convex and bounded subset D of $\mathbb{S}$ which contains all terms of the sequences $\{x^k\}$ and $\{x^k + d^k\}$ for all $k \geq \bar{K}_1$. By Lipschitz continuity of ∇g on D and letting $x = x^k$, $y = x^k + d^k$ and $t > 0$ in Proposition 7.1 we have,

$$g(x^k + td^k) \leq g(x^k) + t\langle \nabla g(x^k), d^k \rangle - \frac{1}{2}L_g t^2 ||d^k||^2, \tag{36}$$

where L_g is the Lipschitz constant for ∇g on D. Using (26) in (36) we obtain

$$g(x^k + td^k) - g(x^k) \leq -(\eta - \frac{1}{2}L_g t)t||d^k||^2$$

hence,

$$g(x^k + td^k) - g(x^k) \le -\sigma t ||d^k||^2$$

holds for all t satisfying $0 \le t \le \min\{1; 2(\eta - \sigma)/L_g\}$, provided that $\sigma < \eta = \mu$. Let,

$$\alpha' = \min\{\beta; 2\beta(\mu - \sigma)/L_g\}$$

Thus by choosing $\alpha_k > \alpha' > 0$, the Armijo step length rule is satisfied for all $k \ge \bar{K}_1$. Hence the sequence α_k is bounded below.

Hence, by boundedness of α_k and (35) it is easy to show that,

$$\lim_{k \to \infty} ||d^k|| = 0. \tag{37}$$

Claim 2: $\lim_{k \to \infty} \Omega(x^k, W(x^k)) = 0$.

For any $x \in \mathbb{S}$, we have $\Omega(x, x) = 0$. Using the Lipschitz continuity of Ω in the second variable we can write,

$$\Omega(x^k, W(x^k)) = \Omega(x^k, W(x^k)) - \Omega(x^k, x^k) \le L_\Omega ||W(x^k) - x^k||$$

hence,

$$\Omega(x^k, W(x^k)) \le L_\Omega ||d^k|| \tag{38}$$

taking limit on both side and using (37)

$$\lim_{k \to \infty} \Omega(x^k, W(x^k)) = 0 \tag{39}$$

By definition,

$$\begin{aligned} g(x^k) &= \langle F(x^k), x^k - W(x^k)\rangle - \Omega(x^k, W(x^k)) \\ &= \langle F(x^k), d^k)\rangle - \Omega(x^k, W(x^k)) \end{aligned} \tag{40}$$

taking limit on both side of (40) and using (39) we obtain,

$$\lim_{k \to \infty} g(x^k) = 0$$

(iii) $\lim_{k \to \infty} x^k = x^*$, where x^* is the unique solution to the VIP (1).

Since F is strongly monotone and continuous, VIP (1) has a unique solution, say x^*. By Theorem 4.2,

$$g(x_k) \ge \alpha ||x_k - x^*||^2,$$

holds for some fixed $\alpha > 0$. Taking limit,

$$\lim_{k \to \infty} x^k = x^*$$

this completes the proof. $\square$

8. Conclusion

In our present chapter, we introduce a general class of gap function for the variational inequality problem in a Hilbert Space. We prove some excellent properties of the proposed gap function and based on those properties we develop a very flexible descent method to solve the VIP. The liberty of constructing the function Ω in the definition of gap function makes the method well adaptive to the problem environment. The global convergence of the proposed iterative algorithm is proved under the strong monotonicity assumption on the main operator (F). It is, however, worth mentioning, that the strong monotonicity assumption is satisfied very often in the problem of Mathematical physics; e.g, see [19].

REFERENCES

[1] *Y. Yao, M. Postolache, J.C. Yao*, An iterative algorithm for solving generalized varia-tional inequalities and fixed points problems, Mathematics, **7**(1)(2019), Art. No. 61.

[2] *Y. Yao, M. Postolache, J.C. Yao*, Iterative algorithms for generalized variational inequa-lities, U. Politeh. Buch. Ser. A, **81**(2)(2019), 3-16.

[3] *Pitea, Ariana*, Results on some classes of vector variational inequalities, Nonlinear Convex Anal., **19**(3)(2018), 417–432.

[4] *Y. Yao, M. Postolache, Y.C. Liou, Z. Yao*, Construction algorithms for a class of monotone variational inequalities, Optim. Lett., **10**(7)(2016), 1519-1528.

[5] *G. Harker, T. Patrick and Pang, Jong-Shi*, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, Math. Program., **48**(1991), 161-220.

[6] *Y. Yao, M. Postolache*, Iterative methods for pseudomonotone variational inequalities and fixed point problems, J. Optim. Theory Appl., **155**(1)(2016), 273-287.

[7] *Auslender, Alfred*, Optimisation: méthodes numériques, Masson (1976).

[8] *Auchmuty, Giles*, Variational principles for variational inequalities, Numer. Funct. Anal. Optim., **10**(1999), 863-874.

[9] *Fukushima, Masao*, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, Math. Program., **53**(1992), 99-110.

[10] *Patriksson, Michael*, Nonlinear programming and variational inequality problems: a unified approach,Springer Science & Business Media, **23**(2013).

[11] *Konnov, Igor V and Kum, Sangho and Lee, Gue Myung*, On convergence of descent methods for variational inequalities in a Hilbert space, Math. Methods Oper. Res., **55**(2002), 371-382.

[12] *Zhu, Dao Li and Marcotte, Patrice*, Convergence properties of feasible descent methods for solving variational inequalities in Banach spaces, Comput. Optim. Appl., **10**(1998), 35-49.

[13] *Noor, M Aslam*, General algorithm for variational inequalities, J. Optim. Theory Appl., **73**(1992), No. 2, 409-413.

[14] *Zhu, Dao Li and Marcotte, Patrice*, An extended descent framework for variational inequalities, J. Optim. Theory Appl. **80**(1994), No. 2, 349-366.

[15] *Taji, Kouichi and Fukushima, Masao and Ibaraki, Toshihide*, A globally convergent Newton method for solving strongly monotone variational inequalities, Math. Program. **58**(1993), No. 1, 369-383.

[16] *Fukushima, Masao*, Merit functions for variational inequality and complementarity problems, Nonlinear Optimization and Applications, **58**(1993), No. 1, 369-383.

[17] *Peng, Ji-Ming*, Equivalence of variational inequality problems to unconstrained minimization, Math. Program., **78**(1997), No. 3, 347-355.

[18] *Yamashita, Nobuo and Taji, Kouichi and Fukushima, Masao*, Unconstrained optimization reformulations of variational inequality problems, J. Optim. Theory Appl. , **92**(1997), No. 3, 439-456.

[19] *Kinderlehrer, David and Stampacchia, Guido*, An introduction to variational inequalities and their applications, SIAM (2000).

[20] *C.L. Li, M. Postolache, Z.F. Jia*, Weighted method for uncertain nonlinear variational inequality problems, Mathematics (submitted).

[21] *Y. Yao, M. Postolache, J.C. Yao*, Strong convergence of an extragradient algorithm for variational inequality and fixed point problems. U Politeh Buch Ser A (in printing).